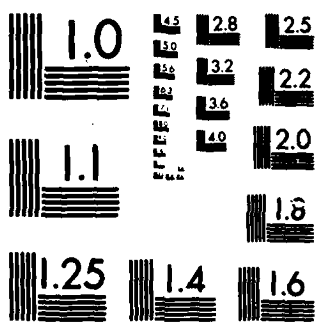
AD-A163 685 GENERALIZING OVER CONDITIONS BY COMBINING THE MULTITRAIT MULTIMETHOD MATR..(U) COLORADO UNIV AT BOULDER CENTER FOR RESEARCH ON JUDGMENT AND P.. 1/1

UNCLASSIFIED K R HAMMOND ET AL. JAN 86 CRJP-255A F/G 5/10 NL

END



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

GENERALIZING OVER CONDITIONS BY COMBINING THE MULTITRAIT MULTIMETHOD MATRIX AND THE REPRESENTATIVE DESIGN OF EXPERIMENTS

Kenneth R. Hammond, Robert M. Hamm and
Janet Grassia

12

GENERALIZING OVER CONDITIONS BY COMBINING THE MULTITRAIT MULTIMETHOD MATRIX AND THE REPRESENTATIVE DESIGN OF EXPERIMENTS

Kenneth R. Hammond, Robert M. Hamm and

Janet Grassia

Center for Research on Judgment and Policy

Institute of Cognitive Science

University of Colorado

Boulder, Colorado 80309

Report No. 255A

January 1986

This research was supported by the Engineering Psychology Programs, Office of Naval Research, Contract N00014-81-C-0591, Work Unit Number NR 197-073 and by BRSG Grant #RR07013-14 awarded by the Biomedical Research Support Program, Division of Research Resources, NIH. Center for Research on Judgment and Policy, Institute of Cognitive Science, University of Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government. Approved for public release: distribution unlimited.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE — READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: CRJP 255A
2. GOVT ACCESSION NO.: ADA 163 685
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Generalizing over Conditions by Combining the Multitrait-Multimethod Matrix and the Representative Design of Experiments
5. TYPE OF REPORT & PERIOD COVERED:
6. PERFORMING ORG. REPORT NUMBER:
7. AUTHOR(s): Kenneth R. Hammond, Robert M. Hamm, & Janet Grassia
8. CONTRACT OR GRANT NUMBER(s): N00014-81-C-0591
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Center for Research on Judgment and Policy, Campus Box 344, University of Colorado, Boulder, CO 80309
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 197-073
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, 800 North Quincy Street, Arlington, VA 22217
12. REPORT DATE: January 1986
13. NUMBER OF PAGES: 64
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report):
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release: Distribution unlimited
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): multitrait-multimethod, representative design, expert judgment, generalization, methodology.
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

Analytical methods should be substituted for the current largely intuitive methods for generalizing results over conditions. Toward that end we present a methodology that combines Campbell and Fiske's (1959) multitrait-multimethod matrix and Brunswik's (1956) representative design of experiments. A description of a study of expert judgment and a critique of current practices illustrate the methodology.

DD FORM 1473, 1 JAN 73 — EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Doubts about the generality of results produced by psychological research have been expressed with increasing frequency since Koch (1959) observed, after a monumental review of scientific psychology in 1959, that there is "a stubborn refusal of psychological findings to yield to empirical generalization" (pp. 729-788). Brunswik (1952, 1956), Campbell and Stanley (1966), Cronbach (1975), Epstein (1979, 1980), Einhorn and Hogarth (1981), Greenwald (1975, 1976), Hammond (1966), Meehl (1978) and Simon (1979) among others, have also called attention to this situation. Jenkins (1974), warned that "a whole theory of an experiment can be elaborated without contributing in an important way to the science because the situation is artificial and nonrepresentative" [italics added] (p. 794). Tulving (1979) makes the startling observation that "after one hundred years of laboratory-based study of memory, we still do not seem to possess any concepts that the majority of workers would consider necessary or important" (p. 27). Nor is it unusual for reviewers of a body of literature to find, as Hastie and Park (in press) do, that over 50 studies have been carried out on a given topic without yielding a definite conclusion. Meehl (1978), summarized the consequences of the failure to develop generalizable findings by saying:

> There is a period of enthusiasm about a new theory, a period of attempted application to several fact domains, a period of disillusionment as the negative data come in, a growing bafflement about inconsistent and unreplicable empirical results, multiple resort to ad hoc excuses, and then finally

people just lose interest in the thing and pursue other endeavors (p. 807).

It is our view that this situation is caused by the lack of an analytical means for generalizing and thus aggregating results. Consequently, aggregation rests largely on researchers' intuitive judgments of what constitutes generality of results over conditions.

In an effort to develop an analytical methodology, and thus contribute to the development of a cumulative science, we build upon two previous methodological suggestions, one from the field of individual differences (the multitrait-multimethod (MTMM) matrix introduced by Campbell and Fiske, 1959), and one from the field of experimental psychology (the representative design of experiments introduced by Brunswik, 1956). Data from a study of experts who made judgments of three concepts by three different methods provided a unique opportunity not only to make use of each of these suggestions but to combine and extend them.

Research Context

The purpose of this study was to examine the relative efficacy of intuitive, quasi-rational and analytical cognition. Twenty engineers judged the aesthetic value, safety, and capacity of 40 highways under three cognitive conditions. Each engineer's judgments were studied in each cell of the diagram presented in Figure 1. Intuition was induced by requiring each expert to judge each concept (aesthetics, safety, capacity) from film strips of one- to three-mile segments of each of the 40 highways. Quasirationality was induced by requiring each expert to judge each concept from bar graphs that presented the values of nine attributes for each highway. Analytical cognition was induced by requiring each engineer to construct a mathematical formula for each concept. An empirical criterion was available for each concept, and thus it was possible to evaluate the accuracy of each expert's judgment in relation to each concept. The criterion for the aesthetic value of each highway was the mean judgment of 91 citizens who judged the highway segments by rating the film strips, or by rating or ranking single frames from the film strips. The criterion for safety was the accident rate for each highway segment averaged over 7 years. The criterion for capacity was the figure calculated by using the procedure from the Highway Capacity Manual 1965 (Highway Research Board, 1965). Each expert devoted roughly 20 hours to the nine sessions, each of which was separated by two-week intervals. (See Hammond, Hamm, Grassia, & Pearson, 1984, for details.)

Insert Figure 1 about here

Plan of the Article

In what follows we first present a description of the Campbell/Fiske MTMM matrix. Second, we extend the MTMM matrix to include a "coherence validity" matrix. Third, we indicate how a "performance validity" matrix incorporates the essential element of Brunswik's representative design of experiments. Fourth, we show how the complementarity of the two matrices leads to the development of a measure of competence. Fifth, we illustrate how the use of this methodology provides an analytical means for evaluating the results of experiments. Sixth, we use these matrices to illustrate the analytical incompleteness and overgeneralization inherent in conventional methods of accumulating results.

Campbell and Fiske's Multitrait-Multimethod (MTMM) Matrix

In their 1959 article, Campbell and Fiske convincingly demonstrated the faults of the conventional single-concept single-operation methodology in the field of individual differences. They showed that results of studies in this field were more likely to be determined by the methods employed than by the traits hypothesized to account for the results. Although they also showed that the failure to separate the effects of method from the effects of concept can be avoided by use of the MTMM matrix, there has been little change in conventional research methodology; current research in this area still fails to systematically separate concept and method (see, for example, Pervin, 1985).

The problem is not that Campbell and Fiske's work has gone unrecognized. It has become a milestone in the methodological literature of psychology, and by 1983 had been cited over 1000 times. Yet in spite of the potential of the MTMM matrix for breaking the grip of a simpleminded operationism on psychological research, the method is for the most part simply not used. Presumably researchers have avoided it for tactical reasons, since it introduces conceptual complexity (which concepts and which methods should be compared?) and requires considerable additional labor and apparatus within a single study (at least two methods and two concepts must be tested). Or perhaps researchers are generally unaware of the ephemeral character of results produced by single-concept single-method operationism. Whatever the reason, among tens of thousands of studies of individual differences, Turner (cited in Fiske, 1982) found only 70 published matrices between 1967 and 1980 (see Fiske, 1982, for a general review). So far as we can ascertain, the MTMM approach is never used in experimental psychology. (Campbell and Fiske cite one exception, an experimental study that employed the MTMM matrix to examine individual differences. In that study also, "the highest correlations are found among different constructs from the same method, showing the dominance of apparatus or method factors so typical of the whole field of individual differences" (1959, p. 86).)

The MTMM matrix, presented in Table 1, is developed from a set of test scores taken from a group of subjects (Campbell & Fiske, 1959). The scores for each subject are correlated over several traits and methods. This illustration is based on data from "three different traits, each measured by three methods, generating nine separate variables" (1959, p. 82). The reliabilities are indicated in parentheses in the main diagonal. The convergent validity coefficients (monotrait-heteromethod) that are derived from measuring the same trait by different methods are shown in the lower diagonals (e.g., .57, .57, and .46 et seq.). A heterotrait-monomethod triangle lies below the main reliability diagonal and a heterotrait-heteromethod triangle lies to either side of each validity diagonal.

Insert Table 1 about here

Campbell and Fiske (1959) note that "a validity value for a variable should be higher than the correlations obtained between that variable and any other variables having neither trait nor method in common. This requirement may seem so minimal and so obvious as to not need stating, yet an inspection of the literature shows that it is frequently not met, and may not be met even when the validity coefficients are of substantial size" (pp. 82-83). Thus, Campbell and Fiske (1959) introduce not only the concept of convergent validity but discriminant validity. A trait should not only be measured by results from different methods which converge upon it, but also discriminated from its rivals.

The value of this methodology has been widely recognized (see, e.g., Brewer & Collins, 1981; Fiske, 1982), and its application will yield definite and useful conclusions regarding the validity of psychological traits. (See Widaman, 1985, who reviews criticisms of the MTMM methodology as originally proposed and offers a rigorous procedure for evaluating MTMM matrices; see Schmitt, Coyle, & Saari, 1977; Kenny, 1979; Marsh & Hocevar, 1983, for the use of confirmatory factor analytic methods to analyze the MTMM matrix. See Schmitt, 1978, for the use of path analysis for evaluating a MTMM matrix; see Farh, Hoffman, & Hegarty, 1984, for a recent application.) The results from such a matrix will have populational and task generality insofar as the trait domain, the apparatus/method domain and the subject domain have been adequately sampled. The results therefore establish the construct validity of the traits investigated, separate from the methods used, within the restraints chosen by the investigator.

Extension of the Campbell/Fiske Approach

In this research context we extend Campbell and Fiske's MTMM matrix from the evaluation of the construct validity of certain (a) traits within the study of (b) individual differences based on (c) group data to (a) the construct validity of judgments of concepts in a coherence validity matrix (b) a performance validity matrix that incorporates criterion measures (and their intra-ecological correlations) for each concept, and (c) the behavior of the individual rather than of the group, although group data can be analyzed as well. (See Hammond, McClelland, & Mumpower, 1980, pp. 115-127 on the advantages of single-subject analysis; also Meehl, 1978, and Serlin & Lapsley, 1985, on the problematic nature of conventional between-group comparisons.)

The coherence validity matrix. By extending the multitrait-multimethod procedure to evaluate the construct validity of an individual's judgments of concepts (as in the present study of highway engineers) we can determine whether the individual has indeed mastered each concept and is able (a) to use it across different methods (convergent validity) and (b) to differentiate it from other concepts (discriminant validity). The analysis of the construct validity of an individual's judgments by means of such a multiconcept-multimethod matrix is analogous to the analysis of the traditional Campbell/Fiske

MTMM matrix, with "concepts" substituted for "traits" in Table 1. Because the coherence of the individual's judgments of concepts can be ascertained from a multiconcept-multimethod matrix, it is called a "coherence validity matrix." Of course, the coherence matrix can be applied to individual behavior other than judgments. Problem-solving behavior, memory, and similar cognitive functions as well as psychomotor functions can also be evaluated by means of this matrix.

The use of the coherence validity matrix. Data for a coherence validity matrix based on our study of highway engineers are presented in Table 2. The data for the matrix were generated from the means of the 20 engineers' judgments for each of 40 highways presented to them for each concept-method pair. (Note: Fisher's z-transformation was used throughout this study in the calculation of mean values.) Thus, the matrix illustrates the particulars of the behavior of an artificial engineer constructed from the mean judgments of this group. Data from the artificial engineer are presented mainly to illustrate the use of the method; no inferences can be drawn from the matrix in Table 2 to a matrix generated for any one engineer. Illustrations of individual matrices are provided in Hammond, Hamm, and Grassia (1984).

Insert Table 2 about here

Each of the descriptions of the matrix presented by Campbell and Fiske (1959) apply to the matrix in Table 2. The three validity diagonals contain values that provide evidence regarding convergent validity, and the heteroconcept triangles adjacent to them provide evidence regarding discriminant validity.

Brunswik's Representative Design

Brunswik's (1943, 1952, 1956) argument that generalization over conditions requires the representation of ecological conditions in the design of experiments must also be considered a milestone in the methodological literature of psychology (see Hammond & Wascoe, 1980, for some examples of the use of representative designs). Brunswik's work has also been cited over 1000 times, yet representative designs are seldom employed. Presumably, the same reasons that lead students of individual differences to forgo the use of the MTMM matrix also lead experimental psychologists to forgo the use of representative design: both are more difficult and *time-consuming to execute than* standard experiments in which the central effort--clear separation of one condition from another--is carried out without regard to the arrangement of conditions in the organism's natural habitat. As Brunswik (1956) observed, however: "generalization of results concerning the relative weights of the variables involved must remain limited unless at least the range, but better also the distribution of the 'levels of strength' employed for each variables, has been made representative of a carefully defined universe of conditions (p. 55). "Carefully defining a universe of conditions" and representing them is far more difficult than merely separating them according to design requirements, yet that is what generalization requires.

The Performance Validity Matrix

Table 3 presents a further extension of the Campbell/Fiske MTMM matrix. By linking the nine method-concept conditions with criterion measures for the three concepts, it becomes possible to evaluate the performance of each engineer across different concepts and different methods. The three validity diagonals contain monoconcept correlations between each set of judgments (one for each method) and the criterion for the same concept. For example, in the upper left-hand corner of Table 3, .855 is the correlation between the artificial engineer's judgments, under the film-strip method of the aesthetic value of each highway, and the aesthetic criterion value for each highway. The triangles consist of heteroconcept correlations between the judgments made in each concept-method condition and the criterion for a different concept. Thus, the number .016 (just below .855) represents the correlation between the judgments of aesthetics for each highway and the safety criterion for each highway. Because this multiconcept-multimethod matrix can be used to evaluate an individual's performance when his/her judgments are compared with a criterion measure, we call it a "performance validity matrix."

Insert Table 3 about here

The coefficients in the performance validity matrix in Table 3 are different from those in the coherence validity matrix in that each correlation in the performance validity matrix is between judgments and measures of a criterion rather than between two sets of responses. Aside from this very important difference, the interpretation of the coefficients with respect to the questions of convergent and discriminant validity is quite similar. As in the coherence validity matrix, correlations in the validity diagonal that are sufficiently large are evidence of convergent validity. In Table 3 the coefficients in the diagonals within each method block show the convergent validity of the judgment of each concept by that method. Comparison of the average of these diagonal values across the three concepts indicate the relative external convergent validity of each method. The heteroconcept triangles consist of the correlations of an expert's judgments of one concept (by a particular method) with the criterion measure of a different concept. Evidence of discriminant validity exists when a value in a validity diagonal is higher than the values lying in its column and row in the heteroconcept triangles. (Precise tests of discriminant validity for group data are described in Widaman, 1985; see also Hammond, Hamm, & Grassia, 1984.)

The Use of the Performance Validity Matrix

Convergent validity of concepts and methods. In Table 3, the validity correlation coefficient for the artificial engineer's aesthetics judgments made by the film strip method is .855, by the bar graph method is .945, and by the formula method is .951, thus producing a mean external convergent validity value across all three methods of .926 for aesthetic judgments. Averaging validity correlations pertaining to safety from the three method boxes, the mean convergent validity (Pearson r) is .568; similarly, averaging the judgment-criterion correlations for capacity produces a mean convergent validity value of .530. In short, the data suggest that, irrespective of the method used, the artificial engineer judged highway aesthetics more accurately than highway safety or capacity, and judged safety and capacity with equal accuracy.

Convergent validity of methods. A measure of the performance convergent validity for each method may be calculated by averaging the judgment-criterion correlations within each of the diagonals (.855, .702, .291; .945, .683, .833; and .951, .226, .266), thus obtaining performance validities for each method (.672, .854, and .654). These results suggest that the artificial engineer judged these three concepts most accurately using the bar graph method. Finally, the mean of the latter three coefficients is .742. This measure is informative because it may be used to compare one group of experts with another, to compare one individual with another (in the case when a matrix is constructed for each individual), or to evaluate the effect of a change in condition in either case. Moreover, the referential domain of this measure is clear; it is general over the three methods and three concepts employed in the study, as well as the group of engineers selected.

Discriminant validity. The performance validity matrix provides several measures of discriminant validity. One is analogous to Campbell and Fiske's (1959) first method for calculating discriminant validity from the multitrait-multimethod validity matrix (see above). Other measures, however, have the advantage of taking into account the intra-ecological correlations.

Measuring performance discriminant validity using a procedure analogous to Campbell and Fiske's. The discriminant validity of a specific concept-method unit can be determined using the first measure, by comparing its validity coefficient with the heteroconcept coefficients that include the concept of interest. Thus, for example, in Table 3 the magnitude of the validity correlation for the aesthetics judgment made by the film-strip method (.855) can be compared with the magnitude of the correlations between the aesthetics criterion and the safety and capacity judgments (-.362 and -.473, respectively), as well as with the size of the correlations between aesthetics judgments and the safety and capacity criteria (.016 and -.172, respectively). Although subjective appraisals of such comparisons may suffice, objective comparisons may be provided by subtracting the mean of the four heteroconcept correlations from the monoconcept correlation of interest (after appropriate z-transformations and sign changes; see discussion of Hypothesis 2, below).

Averaging the performance discriminant validity measures for the artificial engineer's aesthetics judgments across film strip (.598), bar graph (.611), and formula (.736) methods produces a measure of .653 for the artificial engineer's performance discriminant validity with regard to aesthetics; similarly, for safety the performance discriminant validity is .178 and for capacity, .115. Analogous procedures produce measures of the performance discriminant validity of each method.

Measuring discriminant validity with reference to intra-ecological correlations. The correlation between the criterion measures of the concepts provides a standard against which to compare the heteroconcept correlations between the expert's judgment of one concept and the criterion measure of a different concept. For example, since the correlation between aesthetics criteria and safety criteria is -.275, then it is appropriate for an engineer's judgments of aesthetics to be correlated -.275 with safety. Similarly, if since the correlation between two criterion measures is low (as for safety and capacity, .180), then the heteroconcept correlations should also be low. (See left-hand side of Table 3 for intra-ecological correlations of criteria.) In short, the observed correlations between judgments of aesthetics, safety and capacity for an engineer are not to be compared to a standard of zero (an arbitrary demand for complete independence regardless of task conditions) but to a standard that is representative of task conditions, if we are properly to evaluate the discriminant validity of the judgments of these concepts with these methods.

To "untie" these variables, in other words to force them to be orthogonal to one another, is (a) to invite the engineer to judge an unrepresentative set of conditions and thus (b) to extrapolate the results obtained illegitimately from irrelevant conditions to the relevant ones. These two tactics have an embarrassingly long history in psychology; they are customarily explained away by arguments that "this is the best we can do" and/or "it doesn't matter, anyway." Neither argument is correct, but neither is necessary; the performance validity form of the multiconcept-multimethod matrix makes it possible to evaluate the competence of experts (or other subjects) in relation to the task conditions to which their judgments are to be applied.

Discriminant validity of concept pairs. A performance validity measure which relates judgments to intra-ecological correlations is therefore preferable to the first performance discriminant validity measure, described above. The performance discriminant validity of a pair of concepts (e.g., aesthetics and safety) can be determined with the second performance validity measure as follows. Each of the correlations in Table 3 involving aesthetics and safety (-.362, .016, -.479, -.233, -.226 and -.313) and the intra-ecological correlation between the criterion measures of aesthetics and safety (-.275) are z-transformed. The difference between each aesthetics-safety correlation and the intra-ecological correlation is computed. The absolute values of these differences are averaged; and the mean, .129 for the aesthetics-safety example, is an index of performance discriminant validity. The corresponding index for aesthetics and capacity is .121, and for safety and capacity, .302. This indicates that the safety and capacity concepts are discriminated least accurately.

Complementarity of the Coherence and Performance Validity Matrices Leads to a Measure of Competence

The distinction between coherence and performance is intended to parallel the traditional distinction between the coherence and correspondence theories of truth (see, e.g., White, 1967, and Prior, 1967). The coherence theory focuses on the extent to which statements of facts or judgments put forward <u>cohere</u>, or "hang together" with one another; that is, are related by logical implication. Like the coherence theory of truth, the coherence validity matrix demands logical rather than external justification. Although the coherence matrix does include empirical, factual material, no reference to empirical criteria <u>outside the matrix itself</u> is required to establish the construct validity of a set of psychological concepts. All that is required is that a logical criterion be met, namely, that convergent validities should be high and discriminant validities should be low.

The correspondence theory of truth, on the other hand, is concerned with the extent to which our beliefs about the world correspond to independently determined facts. Therefore an independent measure of the concepts in question is required in order to test the correspondence between what a theory predicts and what exists. The performance validity matrix thus parallels the correspondence theory of truth in that it demands the evaluation of the performance of a theory; it demands the evaluation of the empirical correspondence between psychological concepts and some independent measure of the concepts. (See Einhorn & Hogarth, 1982, for a similar treatment of the components of expert judgment in which "truth," our "coherence," combined with "accuracy," our "performance," constitute "justifiability".)

Because both matrices can be developed for a single subject (and aggregated for group analyses), it is possible to <u>combine</u> the results from each matrix into a single measure to provide a higher order indicator of each individual's judgment that we shall call "competence" (see also McClelland, 1973). Since we derive the measure of competence from measures of coherence and performance that are based on variations in both method and concept, our derivation copes directly with the problem of generalization. In the present case, for example, the conclusions about an expert's coherence and performance, and thus competence, are clearly based on, and thus limited to, his behavior over the three methods and three concepts employed in the study. (J. R. Kirwan, personal communication, December, 1985, has innovatively applied this method to physicians' judgments.)

Indices of Coherence, Performance and Competence

Individual differences among engineers can be studied using numerical measures of convergent and discriminant validity derived from both the coherence and performance validity matrices. Procedures for evaluating validity can be converted into numerical measures by adding (or subtracting, as appropriate) the correlations in all the relevant cells in the matrix. These measures can be combined into indices of overall coherence validity, which indicates the coherence of the engineer's judgments; the corresponding index of performance validity indicates the correspondence between the engineer's judgments and reality. And the mean of these two indices provides a measure of the engineer's overall competence. Each index can be produced at different levels of aggregation (e.g., for each concept or for each pair of methods), thus allowing numerical comparisons among these indices at each level. (The formulas for producing these indices can be found in Appendix A.)

The treatment of coherence and performance as separate cognitive functions thus will allow us to examine empirically questions of fundamental importance to both philosophers and psychologists. For example,

1. Does high competence always imply both high coherence and high performance? Are coherence and performance always functionally interdependent? Can one exist in the absence of the other?

2. Should the measures of coherence and performance be equally weighted components of competence? Currently, different approaches to the study of cognition weight these components of competence differently. For example, students of artificial intelligence and problem solving weight the experts' coherence (and the coherence of the computer program that simulates the expert) very highly while virtually ignoring performance. Students of judgment and decision making do the opposite (Hammond et al., 1980). These two fields of research are apparently investigating complementary aspects of competence among experts.

A fairly high correlation (.60) was found between coherence and performance among the experts in the example used here.

Application to Experimental Psychology

In basic research, the results of experiments are expected to be general, that is, not contingent upon the use of a single method. In addition, it is expected that the experiments will show that the concepts employed to describe the results, and not others, are indeed responsible for the results. We indicate how the construction of the coherence and performance validity matrices provides a systematic approach to the testing of hypotheses regarding these desiderata.

Coherence Validity Matrix

The first requirement is that a coherence validity matrix, similar to those in Tables 2 and 3 above, be produced for each engineer, and that convergent and discriminant validities be determined for each.

Convergent validity. The coherence convergent validity measure (monoconcept heteromethod correlations between judgments of the same concept using different methods) can be used to test the primary hypothesis of interest, thus:

> H1: Each theoretical concept has empirical meaning independent of a specific method, i.e., there is *convergent validity for each concept across* methods and within an appropriate sample of subjects.

Hypothesis 1 was tested by asking whether, for each subject, judgments of a concept covary, independently of the methods used to make the judgments. For example, for the artificial engineer (Table 2) the correlation between the film strip and bar graph methods for the aesthetics concept is .890; for the film strip and formula methods, .864; and for the bar graph and formula methods, .985. The overall convergent validity for aesthetics is the mean of these correlations (z-transformed), .938, which is significant at $p < .001$. In addition to the matrix for the artificial engineer, a matrix was developed for each of the 20 engineers individually, and this procedure was carried out for each of the three concepts. All 20 engineers had significant positive convergent validities for aesthetics, 16 for safety, and 17 for capacity. Hence we conclude that each of the three concepts is capable of being measured by appropriate subjects independently of the method used; generality has been achieved over three methods. The generality of results regarding more specific hypotheses may also be addressed. For examples see Hammond, Hamm, and Grassia (1984).

Discriminant validity. Convergent validity informs us about the covariance of judgments across methods, and thus about the status of a concept independent of the method used to measure it. In addition, however, we need to know whether the concept is discriminable from other proposed theoretical entities. Campbell and Fiske (1959) gave first priority to this test; for although many people would think it "so minimal and obvious as not to need stating," (p. 82) they observed that it often fails to be true. The coherence discriminant validity analysis employed in the examples below compares monoconcept heteromethod correlations to heteroconcept heteromethod correlations and thus allows an evaluation of the discriminability of each concept. It also permits a more detailed investigation, thus:

H2: All pairs of concepts are equally discriminable.

This hypothesis was tested by calculating an index for each concept pair for each engineer, and looking for evidence of any concept being more, or less, discriminable than the others, for a statistically significant number of engineers. To illustrate the calculation of the index for the aesthetic and safety concepts, for the artificial engineer of Table 2, we compare the correlations from the validity (monoconcept heteromethod) diagonals that involve either aesthetics (.890, .864, .985) or safety (.713, .393, .422) with the correlations from the heteroconcept heteromethod triangles that involve both concepts (.283, .244, .360, .093, .548, and .209). (The sign on all heteroconcept correlations involving aesthetics was reversed because the intra-ecological correlations between the criterion measures of aesthetics and safety, and of aesthetics and capacity, were negative.) In order to aggregate these comparisons into an index, we subtract the mean of the z-transformations of the second set of correlations (.306) from the mean of the z-transformations of the first set (1.155), which produces an index (.849) of the discriminability of the aesthetics and safety concepts. The corresponding index for aesthetics and capacity is .913; for safety and capacity, -.047. Thus, for the artificial engineer aesthetics and capacity are the easiest concepts to discriminate, and safety and capacity are most difficult to discriminate, a result which carries practical implications.

This index of discriminant validity is calculated for each concept pair from each subject's matrix, and the order among concept pairs is determined. For all 20 engineers, the safety and capacity concepts were least discriminable (Chi-squared = 37.053, $p < .001$). Therefore null Hypothesis 2 is rejected, for the engineers' judgments of safety and capacity are more similar to each other than either is to their judgment of aesthetics.

Performance Validity Analysis

Information beyond coherence validity is necessary for experimental confirmation of hypotheses regarding the phenomena of interest, in this case, judgments of aesthetics, safety and capacity.

Convergent validity. In the present case, measures of performance convergent validity can be based on the correlation between an engineer's judgments of a concept and the criterion measure of that concept. Rather than simply testing hypotheses regarding the performance convergent validity of each concept across methods we test a more informative hypothesis, the relative convergent validity of each concept, thus:

> H3: No concept has higher or lower performance convergent validity than any other.

Hypothesis 3 is tested by averaging the z-transforms of the correlations for each concept across methods, and then comparing the averages for each concept. The aesthetics concept had higher convergent validity than safety or capacity for all 20 engineers (Chi-squared = 37.053,

$p < .001$). Despite, or because of, the counterintuitive nature of this result, it has a claim to our attention for it is general across three methods and stands against two other concepts. Similar questions of performance convergent validity can be addressed to methods. For examples, see Hammond, Hamm, and Grassia (1984).

Discriminant validity. The critical question is whether the concepts in question are discriminable by the subjects. Again we inquire into the relative discriminability of the concepts of interest *with regard to performance*, thus,

H4: All pairs of concepts are equally discriminable.

Hypothesis 4 can be tested by calculating an index of discriminability for each concept pair for each engineer, determining the order of these indices for each engineer, and seeing whether any particular order occurred in a significant number of engineers. The calculations in the first step of this procedure were illustrated above in the discussion of the artificial engineer's performance in discriminating between pairs of concepts (see Table 3). The procedure is carried out for each engineer, producing a measure of how well he discriminates each pair of concepts, over all three methods. The safety and capacity concepts were least accurately discriminated for 18 of the 20 engineers (Chi-squared = 26.4, $p < .001$). Tha is, 18 of the 20 engineers underdiscriminate safety and capacity, a result that is consistent with the results from the coherence validity analysis.

Summary

In this section we have illustrated the application of the multiconcept-multimethod analysis to topics typically of concern to experimental psychologists, namely, testing general propositions regarding behavior. In the examples given above the convergent and discriminant validity of three concepts was tested across three methods within the same study using the same subjects. This was done by using the coherence validity matrix, which is concerned solely with the relations among different judgments of the concepts obtained under different methods, and with the performance validity matrix, which is concerned with the relation between the experts' judgments and the criterion measures of the concepts.

Our illustration highlights the complementarity of these two analyses. We found in both the coherence and performance validity analyses that the aesthetics concept has the highest convergent validity, and that safety and capacity are least discriminable from each other. The performance validity analysis was able to put this last finding in sharper perspective than could the coherence validity analysis alone. It showed that the experts underdiscriminate safety and capacity in comparison with the intercorrelation between the criterion measures of these concepts. The engineers' judgments of these two concepts are most highly correlated, while in fact the criterion measures of these concepts have the lowest intercorrelation. This could have been otherwise; that is, if the criterion measures of safety and capacity had actually been very highly correlated, the engineers might have overdiscriminated them. The performance validity analysis provides the only way to determine which of these possibilities is true.

The difference between this approach to establishing the generality of results from experimental research can best be understood by contrasting it with conventional methods.

Conventional Methods of Generalization

In spite of their many elaborations conventional methods omit the crucial comparisons of convergent and discriminant validity and thus claim cumulative results without sufficient analytical justification. For an example that typifies the largely intuitive method of accumulating results in conventional practice, we consider Anderson's discussion (1985, pp. 110-112) of experimental evidence for differential memory of abstract and concrete knowledge. Anderson cites two experiments, one regarding the retention of perception-based knowledge and one regarding the retention of meaning-based knowledge. He concludes that the results confirm each other and thus justify the generalization that "we remember abstract information, not details" (p. 112). Anderson's method of accumulating results is typical in that it rests on the observation that one experimenter has used one method to test the validity of an hypothesis, a second experimenter has used a different method to test the same hypothesis, and similar results have been obtained, thus general knowledge is claimed. But examination of the experimental results from the standpoint of the MTMM methodology shows why this claim is not justified.

The first experiment cited by Anderson (1985) was reported in an article by Posner (1969) in which the difference in reaction time to an "identity match" and a "name match" of letters was taken as a measure of differential retention of information. Anderson points out that "initially there is a large advantage [i.e., shorter reaction time] for the identity match but after a two-second [inter-stimulus] interval this advantage has almost completely disappeared. This alteration indicates that memory for the initial stimulus is rapidly transformed into an abstract code that does not retain specific visual information" (pp. 110-111).

The second experiment cited was carried out by Anderson (1974). He remarks that his experiment "makes the same point in the verbal domain as Posner's did in the perceptual domain" (p. 111). In Anderson's (1974) study, as in Posner's, there is the possibility of a specific information match and an abstract code match. Anderson's study differs from Posner's in that the match occurs in the context of choosing among the logical implications of critical sentences in a story. It is similar, however, in that reaction time was also used to evaluate differences between response categories that are analogous to Posner's identity match and name match. Differences in reaction time are again observed to be related to a length of time between presentations of the stimuli (inter-stimulus intervals). Anderson (1985) concludes that his results confirm Posner's, and summarizes with the following generalization: "So, it seems that verbal information, like visual information, tends to be short-lived and that after delays we mainly remember abstract information" (p. 112).

But the conclusion is unjustified; the aggregation process is analytically incomplete because neither convergent validity nor discriminant validity was established in either experiment. Moreover, the meaning of the principal concept (memory for abstract versus concrete materials) is exhausted in both cases by the same single operation (reaction time in relation to inter-stimulus interval), thus offering us no evidence that the result is not confined to the reaction time measure.

The incompleteness of conventional practices for aggregating results can be seen when the Posner and Anderson studies are represented in the multiconcept multimethod matrices of coherence and performance (see Table 4). Even if we assume that (a) the abstract-concrete dimension can be separated into two concepts and (b) Posner and Anderson's studies represent two different methods, we find that, of the four required, only one cell related to discriminant validity can be filled in in each study; no evidence of convergent validity (i.e., validity independent of method) can be provided. Separately, then, each study is incomplete.

Insert Table 4 about here

A performance validity matrix cannot be constructed for either study because of the restriction of the measurement of the abstract-concrete dimension in both studies to reaction time. That is, although monoconcept heteromethod correlations can be calculated (see the top row of the performance validity matrix) these results are limited to the single criterion used; no data are available for other measures of performance such as, for example, accuracy of recall, which, after all, is what the content of tne generalization claims.

If, however, we relax conditions further and pretend that each subject participated in all conditions, then we find that a hypothetical matrix would permit the examination of both discriminant and convergent validity in the coherence validity matrix. By this procedure each cell in the matrix is filled (if we further assume that reliabilities were calculated). That is, all the question marks in Table 4 would be removed. But the use of the same criterion measure (reaction time) in both studies makes it impossible to use the hypothetical aggregation to contrast performance with regard to memory for detail or for abstractions even if the same subject participated in all conditions.

Anderson (1985) cites a third study, however, that was carried out by Kolers (1979), that does use a direct criterion--accuracy of recall--for evaluating the retention of pictorial material. Unfortunately for Anderson's generalization, Kolers (1979) found results opposite to those obtained by Posner and Anderson. Anderson states that: "In a series of clever experiments, Kolers (1979) has shown that under appropriate conditions we can retain visual details about the typography of a page of print for months!" (Anderson, 1985, p. 112). Because Kolers used a different criterion of performance for the generalization from that used by Posner and Anderson, we are thus left with two contradictory performance validity matrices. But since each incomplete, we cannot be certain which conclusion should be denied or accepted.

What, then, are we to conclude, other than that sometimes one gets one result, sometimes another? As matters stand, we cannot reconcile the contradictory results. Nor can we decide what experiment to do next without an analytical framework that indicates what information is needed to disconfirm these and other alternative plausible hypotheses. The coherence and performance matrices, however, make the requirements of future studies obvious because they specify which cells must be filled in order to defend the generalization. (See Farell, 1985, for a description of the wide variety of methods and concepts that must be considered in relation to "same"-"different" judgments.)

Our aim, of course, is not to single out for criticism the above studies or Anderson's way of reporting them. Rather, it is our intention to illustrate the largely intuitive, analytically incomplete, conventional method of aggregation and to urge its replacement by the analytical method described here, or by better ones (see, Meehl, 1978, for severe criticism of the conventional methods of cumulating results and a recommendation for improvement). But since psychologists' (and other scientists') cognitive strategies for asserting confirmation and/or generalization of results are largely intuitive they can be studied, and thus described, from the standpoint of judgment and decision theory (Hammond et al., 1980, Einhorn & Hogarth, 1981; Pitz & Sachs, 1984), and discrepencies between scientists' cognitive strategies and normative procedures could be examined with considerable profit, no matter what the results turned out to be. Tversky (1977), Tversky and Gati (1982) and Gati and Tversky (1984) have shown that judgments of similarity and dissimilarity, which, of course, are at the root of scientists' judgments of confirmation and/or generalization, can be analyzed and understood in terms of the relative weights attached to common and distinctive features of various entities. (Other methods are described in Hammond et al., 1980.) Such studies would provide descriptions of scientists' judgment processes, which could then be compared to prescriptive, (normative) means of aggregating results, and thus enable us to discover the nature and extent of the differences between them. But in order to accomplish that goal, a prescriptive, normative methodology such as we have described here must be provided; no other exists at present.[1]

Summary and Discussion

As several psychologists have observed, psychological research lacks the cumulative character critical to the development of a science. In any such circumstance suspicion would arise that the scientific discipline in question is the captive of a flawed theoretical or methodological dogma. Since theories are numerous in psychology, but methodology is uniform throughout graduate schools and journal reviews, dogmatic methodology must be the prime suspect.

We extended and integrated the pioneering efforts of Campbell and Fiske (1959) and Brunswik (1956) in order to replace the current judgment-based method with an analytically based methodology for achieving generalization over conditions as well as subjects. We then presented an example of how generalization over methods and concepts can be obtained by the use of the coherence and performance validity matrices.

The logic of the coherence validity matrix is based on what Feigl called "triangulation in logical space" (Feigl, 1958; see also Campbell & Fiske, 1959, p. 84). From a logical point of view, the methods and concepts selected for study should be completely independent; the "triangulation" should approximate a right triangle as nearly as possible. Thus, Campbell and Fiske (1959) discuss "convergence of the independent methods" and cite Cronbach and Meehl's argument that the use of "diverse criteria give[s] greater weight to the claim of construct validity than do ... predictions of very similar behavior" (Cronbach & Meehl, 1955, p. 295).

Brunswik, however, emphasized the fact that the ecological variables that so often serve as criteria for psychologists' concepts are not independent, i.e., orthogonal to one another, in the organism's natural habitat. At the very least, such independence should not be taken for granted and uncritically made the essential design feature of every experiment. Therefore, from the researcher's point of view, Fiegl's concept of "triangulation in logical space" is not to be seen as a goal, but as a condition that serves didactic purposes, without regard to the demands of specific problems. The proper goal for the researcher (in contrast to the logician) is "triangulation in empirical space," in which the logician's worship of orthogonality is replaced by the researcher's worship of empirical generalization. Informative as the logician's remarks undoubtedly are, the proper goal of basic research is generalization of results. That goal can best be achieved through the use of "representative triangulation," and through the use of a performance validity matrix, in experiments as well as in studies of individual differences. We demonstrated how the conventional method of cumulating results is analytically incomplete, and thus largely intuitive. Therefore, unjustified claims of generality are to be expected.

But representative triangulation will require a substantial shift in methodology that takes cognizance of (a) the congruence between the strategy of conventional experimental designs and the aims of applied research, and (b) the incongruence between conventional experimental designs and the aims of basic research. Conventional research does not demand that claims of generalization over conditions or "treatments" be analytically justified (as our example showed). It does demand that generalization over subjects be justified (as the plethora of statistical tests over populations demonstrates). Justification of generalization over subjects is to be expected in applied research, especially applied agricultural research, which is the source of conventional designs in experimental psychology (cf. Newell & Simon, 1972, who make a similar observation). That is because applied agricultural research is disinterested in generalizing over conditions; once the treatment effects have been established, that is, found to be general over the relevant subjects (plants or animals), the farmer can control future conditions, i.e., apply only the "treatment" that works. Thus, the research user gets the information s/he wants. But controlling conditions outside the experiment is precisely what basic research psychologists cannot do. In lieu of applying the successful treatments as the farmer does, they must assert, by intuitive judgments of similarities and dissimilarities between different laboratory experiments, that the results provide evidence of confirmation and/or generalization. Thus, basic researchers use exactly the wrong strategy, namely, fixed conditions and general subjects, a strategy requiring generalization by judgments about what constitutes confirmation and/or generalization. Although considerable progress in understanding the nature of scientists' "generalization by judgment" might well be achieved by means of the various methods of judgment analysis, basic researchers should employ a strategy appropriate to basic research, together with an analytical method for justifying claims of generality. One such analytical method is described here.

References

Anderson, J. R. (1974). Verbatim and propositional representation of sentences in immediate and long-term memory. Journal of Verbal Learning and Verbal Behavior, 13, 149-162.

Anderson, J. R. (1985). Cognitive psychology and its implications (2nd ed.). New York: Freeman.

Brewer, M., & Collins, B. (Eds.). (1981). Scientific inquiry and the social sciences. San Francisco: Jossey-Bass.

Brunswik, E. (1943). Organismic achievement and environmental probability. Psychological Review, 50, 255-272.

Brunswik, E. (1952). The conceptual framework of psychology. In International encyclopedia of unified science (Vol. 1, No. 10). Chicago: University of Chicago Press.

Brunswik, E. (1956). Perception and the representative design of psychological experiments (2nd ed.). Berkeley: University of California Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validiation by the multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Campbell, D. T., & Stanley, J. (1966). Experimental and quasi-experimental designs for research. Chicago: Rand-McNally

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, 281-302.

Dawes, R. M. (1970). An inequality concerning correlation of composites vs. composites of correlations. Oregon Research Institute, Methodological Note, Vol. 1, No. 1.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. Annual Review of Psychology, 32, 53-88.

Einhorn, H. J., & Hogarth, R. M. (1982). Reply to commentaries. In G. Ungson & D. Braunstein (Eds.), Decision making: An interdisciplinary inquiry (pp. 53-56). Boston: Wadsworth.

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. Journal of Personality and Social Psychology, 37, 1097-1126.

Epstein, S. (1980). The stability of behavior II. Implications for psychological research. American Psychologist, 35, 790-806.

Farell, B. (1985). "Same"-"different" judgments: A review of current controversies in perceptual comparisons. Psychological Bulletin, 98, 419-456.

Farh, J-L., Hoffman, R. C., Hegarty, W. H. (1984). Assessing environmental scanning at the subunit level: A multitrait-multimethod analysis. Journal of the American Institute for Decision Sciences, 15(2), 197-220.

Feigl, H. (1958). The mental and the physical. In H. Feigl, M. Scriven, and G. Maxwell (Eds.), Minnesota studies in the philosophy of science (Vol. 2): Concepts, theories and the mind-body problem (pp. 370-497). Minneapolis: University of Minnesota Press.

Fiske, D. W. (1982). Convergent-discriminant validation in measurements and research strategies. In D. Brinberg & L. H. Kidder (Eds.), Forms of validity in research (pp. 77-92). San Francisco: Jossey-Bass.

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. Cognitive Psychology, 16, 341-370.

Greenwald, A. G. (1975). Significance, nonsignificance, and interpretation of an ESP experiment. Journal of Experimental Social Psychology, 11, 180-191.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use? Psychological Bulletin, 83, 314-320.

Hammond, K. R. (Ed.). (1966). The psychology of Egon Brunswik. New York: Holt, Rinehart, & Winston.

Hammond, K. R., Hamm, R. M., & Grassia, J. (1984). Achieving generality over conditions: Combining the multitrait multimethod matrix and the representative design of experiments (Technical Report No. 255). Boulder: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R., Hamm, R. M., Grassia, J., & Pearson, T. (1984). The relative efficacy of intuitive and analytical cognition: A second direct comparison (Technical Report No. 252). Boulder: University of Colorado, Center for Research on Judgment and Policy.

Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). Human judgment and decision making: Theories, methods, and procedures. New York: Praeger.

Hammond, K. R., & Wascoe, N. E. (Eds.). (1980). Realizations of Brunswik's representative design. San Francisco: Jossey-Bass.

Hastie, R., & Park, B. (in press). The relationship between memory and judgment depends on whether the judgment task is memory-based or online. Psychological Review.

Highway Research Board. (1965). Highway Capacity Manual 1965 (Highway Research Board Special Report No. 87, National Research Council Publication 1328). Washington, DC: National Research Council.

Jenkins, J. (1974). Remember that old theory of memory? Well, forget it! American Psychologist, 29, 785-795.

Kenny, D. A. (1979). Correlation and causality. New York: Wiley.

Koch, S. (1959). Epilogue. In S. Koch (Ed.), Psychology: A study of a science (Vol. 3, pp. 729-788). New York: McGraw-Hill.

Kolers, P. A. (1979). A pattern analyzing basis of recognition. In L. S. Cermak & F. I. M. Craik (Eds.), Levels of processing in human memory (pp. 363-384). Hillsdale, NJ: Erlbaum.

Light, R., & Pillemer, D. (1984). Summing up. Cambridge, MA: Harvard University Press.

Marsh, H. W., & Hocevar, D. (1983). Confirmatory factor analysis of multitrait-multimethod matrices. Journal of Educational Measurement, 20(3), 231-248.

McClelland, D. C. (1973). Testing for competence rather than for "intelligence." American Psychologist, 28, 1-14.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. Journal of Consulting and Clinical Psychology, 46, 806-834.

Newell, A., & Simon, H. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Pervin, L. (1985). Personality: Current controversies, issues and directions. Annual Review of Psychology, 36, 83-114.

Pitz, G., & Sachs, N. (1984). Judgment and decision: Theory and application. Annual Review of Psychology, 35, 139-163.

Posner, M. I. (1969). Abstraction and the process of recognition. The Psychology of Learning and Motivation, 3, 43-100.

Prior, A. N. (1967). Correspondence theory of truth. In P. Edwards (Ed.), The encyclopedia of philosophy: Vol. 2 (pp. 223-232). New York: Macmillan.

Schmitt, N. (1978). Path analysis of multitrait-multimethod matrices. Applied Psychological Measurement, 2, 157-173.

Schmitt, N., Coyle, B. W., & Saari, B. B. (1977). A review and critique of analyses of multitrait-multimethod matrices. Multivariate Behavioral Research, 12, 447-478.

Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good-enough principle. American Psychologist, 40(1), 78-83.

Simon, H. A. (1979). Models of thought. New Haven: Yale University Press.

Tulving, E. (1979). Memory research: What kind of progress? In L. G. Nilsson (Ed.), Perspectives in memory research (pp. 19-34). Hillsdale, NJ: Erlbaum.

Tversky, A. (1977). Features of similarity. Psychological Review, 84, 327-352.

Tversky, A., & Gati, I. (1982). Similarity, separability and the triangle inequality. Psychological Review, 89, 123-154.

White, A. R. (1967). Coherence theory of truth. In P. Edwards (Ed.), The encyclopedia of philosophy: Vol. 2 (pp. 130-132). New York: Macmillan.

Widaman, K. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. Applied Psychological Measurement, 9, 1-26.

APPENDIX A

Procedure for Combining Indices of Validity

The various indices (e.g., of internal discriminant validity, external validity, or overall validity) are produced by taking the mean of the appropriate subindices (e.g., the first measure of internal discriminant validity, or external convergent validity) according to the pattern illustrated in Figure A-1. Each subindex is produced for each engineer by taking the mean of z-transformed correlations, from specific locations in the internal or external validity matrices, or the mean of the differences between such z-transformed correlations, corresponding to the comparisons that were illustrated above with Hypotheses 1-4. Table A-1 displays the formulas for each of the 9 subindices, at each of 6 possible levels of aggregation. For example, the formula for the internal convergent validity index, at the concept level of aggregation, is:

$$\underset{\substack{j,k \\ j \neq k}}{M} \; r_{m_j m_k}$$

This index is calculated for each concept m. It is the mean, over all pairs of methods j and k where j is different from k, of the z-transformations of $r_{m_j m_k}$, which is the correlation between two judgments of concept m, using method j and method k. The correlations for the external validity matrix are (with one exception) of form r_{mn_j}; that is, the correlation between the criterion measure of concept m and the engineer's judgment of concept n using method j. M is used as a "mean" symbol, representing a sum of correlations divided by the number of correlations summed over. The correlations involved in producing all the subindices in this table have been z-transformed.

Insert Figure A-1 and Table A-1 about here

Once the subindices are calculated as in Table A-1, they combined as indicated in Figure A-1. Thus, the mean of the three internal discriminant validity subindices (IDV_1, IDV_2, and IDV_3) is the index for internal discriminant validity (IDV); the mean of IDV and the internal convergent validity index (ICV) is the index for coherence or internal validity (IV); and the mean of IV and the index for performance or external validity (EV) is the index for overall competence. Further discussion can be found in Hammond, Hamm, and Grassia (1984).

Author Note

This research was supported by the Engineering Psychology Programs, Office of Naval Research, Contract N00014-81-C-0591, Work Unit Number NR 197-073 and by BRSG Grant #RR07013-14 awarded by the Biomedical Research Support Program, Division of Research Resources, NIH. Center for Research on Judgment and Policy, Institute of Behavioral Science, University of Colorado. Reproduction in whole or in part is permitted for any purpose of the United States Government. Approved for public release; distribution unlimited. We are indebted to Ray Cooksey for his assistance. We also thank Tamra Pearson and Doreen Victor for their assistance.

Footnote

[1]Meta-analysis (see, for example, Light & Pillemer, 1984) does not meet our criteria for aggregating results of experiments because it does not require a distinction between studies that establish discriminant and convergent validity and those that do not.

Table 1

A Synthetic Multitrait-Multimethod Matrix

		Method 1			Method 2			Method 3		
	Traits	A_1	B_1	C_1	A_2	B_2	C_2	A_3	B_3	C_3
Method 1	A_1	(.89)								
	B_1	.51	(.89)							
	C_1	.38	.37	(.76)						
Method 2	A_2	.57	.22	.09	(.93)					
	B_2	.22	.57	.10	.68	(.94)				
	C_2	.11	.11	.46	.59	.58	(.84)			
Method 3	A_3	.56	.22	.11	.67	.42	.33	(.94)		
	B_3	.23	.57	.12	.43	.66	.34	.67	(.92)	
	C_3	.11	.11	.45	.34	.32	.58	.58	.60	(.85)

Note. From "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix" by D. T. Campbell and D. W. Fiske, 1959, Psychological Bulletin, 56, p. 82. Copyright 1959 by the American Psychological Association. Reprinted by permission.

Table 2

Coherence Validity Multiconcept-Multimethod Matrix for Artificial Engineer

METHOD	Concepts	Film Strip FS: E	Film Strip FS: S	Film Strip FS: C	Bar Graph BG: E	Bar Graph BG: S	Bar Graph BG: C	Formula F: E	Formula F: S	Formula F: C
FS	E	(.957)								
FS	S	.035	(.825)							
FS	C	-.117	.951	(.793)						
BG	E	.890	-.283	-.364	(.996)					
BG	S	-.244	.713	.712	-.491	(.968)				
BG	C	-.051	.706	.591	-.282	.681	(.979)			
F	E	.864	-.360	-.452	.985	-.548	-.337	(.999)		
F	S	-.093	.393	.427	-.209	.422	.452	-.232	(.415)	
F	C	.176	.403	.333	.015	.270	.340	.004	.001	(.315)

Key:

( ) monoconcept monomethod (reliability)

monoconcept heteromethod (validity)

heteroconc‹‚t monomethod

heteroconcept heteromethod

Key:

E : Aesthetics

S : Safety

C : Capacity

Table 3

Performance Validity Multiconcept-Multimethod Matrix for Artificial Engineer

		METHOD								
		Film Strip FS			Bar Graph BG			Formula F		
CRITERIA	Concepts	E	S	C	E	S	C	E	S	C
	E	.855	-.362	-.473	.945	-.479	-.311	.951	-.226	.002
E–S: -.275										
E–C: -.279	S	.016	.702	.683	-.233	.683	.516	-.313	.226	.199
S–C: .180										
	C	-.172	.399	.291	-.298	.437	.833	-.316	.383	.266

Key:

monoconcept (validity)

heteroconcept

Key:

E : Aesthetics

S : Safety

C : Capacity

Table 4

Coherence and Performance Validity Matrices for Posner and Anderson Studies

Coherence Validity Matrix

		Posner (Perception)		Anderson (Verbal)	
		M_1		M_2	
		$C_{1_{con}}$	$C_{2_{abs}}$	$C_{1_{con}}$	$C_{2_{abs}}$
M_1	C_1	$?^a$			
	C_2	DV^b	$?^a$		
M_2	C_1	$?C_e$	$?^d$	$?^a$	
	C_2	$?^d$	$?^e$	DV^c	$?^a$

Performance Validity Matrix

Criterion 1 (RT/ISI)	$PV^f{}_{11}$	$PV^f{}_{12}$	$PV^f{}_{11}$	$PV^f{}_{12}$
Criterion ? (Accuracy)	$?^f$	$?^f$	$?^f$	$?^f$

Key:

a : Reliabilities

b : Discriminant validity provided by Posner study

c : Discriminant validity provided by Anderson study

d : Heteromethod-heteroconcept discriminant validities

e : Convergent validities

f : Performance validities

Table A-1

Formulas for Indices of Internal and External Validity at Various Levels of Aggregation

			For Each Concept-Method Unit (m,j)	For Each Concept (m)	For Each Concept Pair (m,n) m≠n
INTERNAL VALIDITY	Reliability		$r_{m_j m_j}$	$j\ r_{m_j m_j}$	
	Convergent Validity		$\underset{k\neq j}{k}\ r_{m_j m_k}$	$\underset{j\neq k}{j,k}\ r_{m_j m_k}$	
	Discriminant Validity	IDV1	$\underset{k\neq j}{k}\ r_{m_j m_k} - \underset{k\neq j}{k}\ \underset{n\neq m}{n}\ r_{m_j n_k}$	$\underset{j\neq k}{j,k}\ r_{m_j m_k} - \underset{j\neq k}{j,k}\ \underset{n\neq m}{n}\ \frac{r_{m_j n_k} + r_{m_k n_j}}{2}$	$\underset{j\neq k}{j,k}\ \frac{r_{m_j m_k} + r_{n_j n_k}}{2} - \underset{j\neq k}{j,k}\ \frac{r_{m_j n_k} + r_{m_k n_j}}{2}$
		IDV2	$r_{m_j m_j} - \underset{n\neq m}{n}\ r_{m_j n_j}$	$j\ r_{m_j m_j} - j\ \underset{n\neq m}{n}\ r_{m_j n_j}$	$j\ \frac{r_{m_j m_j} + r_{n_j n_j}}{2} - j\ r_{m_j n_j}$
		IDV3	$\underset{k\neq j}{k}\ r_{m_j m_k} - \underset{n\neq m}{n}\ r_{m_j n_j}$	$\underset{j\neq k}{j,k}\ r_{m_j m_k} - j\ \underset{n\neq m}{n}\ r_{m_j n_j}$	$\underset{j\neq k}{j,k}\ \frac{r_{m_j m_k} + r_{n_j n_k}}{2} - j\ r_{m_j n_j}$
EXTERNAL VALIDITY	Convergent Validity		$r_{m m_j}$	$j\ r_{m m_j}$	
	Discriminant Validity	EDV1	$r_{m m_j} - \underset{n\neq m}{n}\ \frac{r_{m n_j} + r_{n m_j}}{2}$	$j\ r_{m m_j} - j\ \underset{n\neq m}{n}\ \frac{r_{m n_j} + r_{n m_j}}{2}$	$j\ \frac{r_{m m_j} + r_{n n_j}}{2} - j\ \frac{r_{m n_j} + r_{n m_j}}{2}$
		EDV2	$\underset{n\neq m}{n}\ \lvert r_{m n_j} - r_{mn} \rvert$	$j\ \underset{n\neq m}{n}\ \frac{\lvert r_{m n_j} - r_{mn}\rvert + \lvert r_{n m_j} - r_{mn}\rvert}{2}$	$j\ \frac{\lvert r_{m n_j} - r_{mn}\rvert + \lvert r_{n m_j} - r_{mn}\rvert}{2}$
		EDV3	$k\ \underset{n\neq m}{n}\ \lvert r_{m_j n_k} - r_{mn} \rvert$	$j,k\ \underset{n\neq m}{n}\ \frac{\lvert r_{m_j n_k} - r_{mn}\rvert + \lvert r_{m_k n_j} - r_{mn}\rvert}{2}$	$j,k\ \frac{\lvert r_{m_j n_k} - r_{mn}\rvert + \lvert r_{m_k n_j} - r_{mn}\rvert}{2}$

Table A-1 (continued)

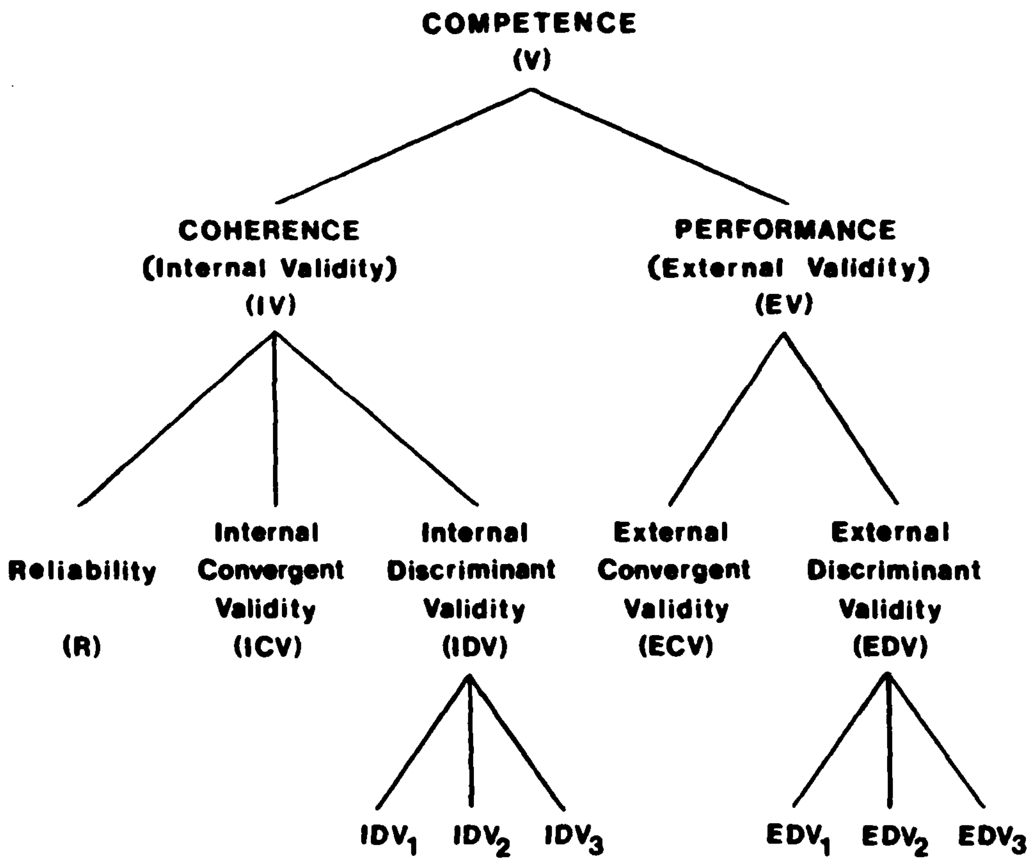
			For Each Method (j)	For Each Method Pair (j,k) j≠k	Overall
INTERNAL VALIDITY	Reliability		$\blacksquare\ r_{m_j m_j}$		$j\ \blacksquare\ r_{m_j m_j}$
INTERNAL VALIDITY	Convergent Validity		$\underset{k\neq j}{k}\ \blacksquare\ r_{m_j m_k}$	$\blacksquare\ r_{m_j m_k}$	$\underset{j\neq k}{j,k}\ \blacksquare\ r_{m_j m_k}$
INTERNAL VALIDITY	Discriminant Validity	IDV1	$\underset{k\neq j}{k}\ \blacksquare\ r_{m_j m_k} - \underset{k\neq j}{k}\ \underset{m\neq n}{m,n}\ \frac{r_{m_j n_k} + r_{m_k n_j}}{2}$	$\blacksquare\ r_{m_j m_k} - \underset{m\neq n}{m,n}\ \frac{r_{m_j n_k} + r_{m_k n_j}}{2}$	$\underset{j\neq k}{j,k}\ \blacksquare\ r_{m_j m_k} - \underset{j\neq k}{j,k}\ \underset{m\neq n}{m,n}\ \frac{r_{m_j n_k} + r_{m_k n_j}}{2}$
INTERNAL VALIDITY	Discriminant Validity	IDV2	$\blacksquare\ r_{m_j m_j} - \underset{m\neq n}{m,n}\ r_{m_j n_j}$		$j\ \blacksquare\ r_{m_j m_j} - j\ \underset{m\neq n}{m,n}\ r_{m_j n_j}$
INTERNAL VALIDITY	Discriminant Validity	IDV3	$\underset{k\neq j}{k}\ \blacksquare\ r_{m_j m_k} - \underset{m\neq n}{m,n}\ r_{m_j n_j}$	$\blacksquare\ r_{m_j m_k} - \underset{m\neq n}{m,n}\ \frac{r_{m_j n_j} + r_{m_k n_k}}{2}$	$\underset{j\neq k}{j,k}\ \blacksquare\ r_{m_j m_k} - j\ \underset{m\neq n}{m,n}\ r_{m_j n_j}$
EXTERNAL VALIDITY	Convergent Validity		$\blacksquare\ r_{m m_j}$		$j\ \blacksquare\ r_{m m_j}$
EXTERNAL VALIDITY	Discriminant Validity	EDV1	$\blacksquare\ r_{m m_j} - \underset{m\neq n}{m,n}\ \frac{r_{m n_j} + r_{n m_j}}{2}$		$j\ \blacksquare\ r_{m m_j} - j\ \underset{n\neq m}{m,n}\ \frac{r_{m n_j} + r_{n m_j}}{2}$
EXTERNAL VALIDITY	Discriminant Validity	EDV2	$\underset{m\neq n}{m,n}\ \frac{\left\lvert r_{m n_j} - r_{mn} \right\rvert + \left\lvert r_{n m_j} - r_{mn} \right\rvert}{2}$		$j\ \underset{m\neq n}{m,n}\ \frac{\left\lvert r_{m n_j} - r_{mn} \right\rvert + \left\lvert r_{n m_j} - r_{mn} \right\rvert}{2}$
EXTERNAL VALIDITY	Discriminant Validity	EDV3	$k\ \underset{m\neq n}{m,n}\ \frac{\left\lvert r_{m_j n_k} - r_{mn} \right\rvert + \left\lvert r_{m_k n_j} - r_{mn} \right\rvert}{2}$	$\underset{m\neq n}{m,n}\ \frac{\left\lvert r_{m_j n_k} - r_{mn} \right\rvert + \left\lvert r_{m_k n_j} - r_{mn} \right\rvert}{2}$	$j,k\ \underset{m\neq n}{m,n}\ \frac{\left\lvert r_{m_j n_k} - r_{mn} \right\rvert + \left\lvert r_{m_k n_j} - r_{mn} \right\rvert}{2}$

Figure Captions

Figure 1. Design of the highway engineers study.

Figure A-1. The structure of indices representing coherence, performance and overall competence.

		METHOD		
		Film Strip (Intuition Inducing) I	Bar Graph (Quasirationality Inducing) Q	Formula (Analysis Inducing) A
CONCEPT	Aesthetics E			
	Safety S			
	Capacity C			


COMPETENCE
(V)
COHERENCE
(Internal Validity)
(IV)
PERFORMANCE
(External Validity)
(EV)
Reliability
(R)
Internal
Convergent
Validity
(ICV)
Internal
Discriminant
Validity
(IDV)
External
Convergent
Validity
(ECV)
External
Discriminant
Validity
(EDV)
IDV_1 IDV_2 IDV_3
EDV_1 EDV_2 EDV_3

OFFICE OF NAVAL RESEARCH

Engineering Psychology Program

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CAPT Paul R. Chatelier
Office of the Deputy Under Secretary
of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D. C. 20301

Department of the Navy

Engineering Psychology Program
Office of Naval Research
Code 442EP
800 North Quincy Street
Arlington, VA 22217-5000 (3 copies)

Aviation & Aerospace Technology
Programs
Code 210
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

CDR. Paul E. Girard
Code 252
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Physiology Neurobiology Program
Office of Naval Research
Code 441NP
800 North Quincy Street
Arlington, VA 22217-5000

CAPT. P. M. Curran
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Mathematics Program
Code 411MA
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Douglas DePriest
Code 411
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5000

Information Sciences Division
Code 433
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

CDR Kent S. Hull
Helicopter/VTOL Human Factors
Office
MS 239-21
NASA/Ames Research Center
Moffett Field, CA 94035

Dr. Carl E. England
Naval Health Research Center
Environmental Physiology
P. O. Box 85122
San Diego, CA 92138

CAPT William M. Houk
Commanding Officer
Naval Medical R&D Command
Bethesda, MD 20814-5055

Department of the Navy

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Mr. R. Lawson
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

CDR James Offutt
Office of the Secretary of Defense
Strategic Defense Initiative Organization
Washington, D.C. 20301-7100

Director
Technical Information Division
Code 2627
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D.C. 23075-5000

Dr. J. S. Lawson, Jr.
4773-C Kahala Avenue
Honolulu, HI 96816

Dr. Neil McAlister
Office of Chief of Naval Operations
Command and Control
OP-094H
Washington, D. C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Dr. A. F. Norcio
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, DC 20375-5000

Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Washington, D. C. 20350

Mr. John Davis
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Department
Code N-71
Naval Training Equipment Center
Orlando, FL 32813

Mr. Norm Beck
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Engineering
Code 441
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. H. Talkington
Engineering & Computer Science
Code 09
Naval Ocean Systems Center
San Diego, CA 92152

Department of the Navy

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Ross Pepper
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D. C. 20380

Dr. L. Chmura
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Letsky
Office of the Chief of Naval Operations (OP-01B7)
Washington, D.C. 20350

Professor Douglas E. Hunter
Defense Intelligence College
Washington, D.C. 20374

CDR C. Hutchins
Code 55
Naval Postgraduate School
Monterey, CA 93940

Human Factors Technology Administration
Office of Naval Technology
Code MAT 0722
800 North Quincy Street
Arlington, VA 22217-5000

CDR Tom Jones
Naval Air Systems Command
Human Factors Programs
NAVAIR 330J
Washington, D.C. 20361

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D. C. 20361

Mr. Philip Andrews
Naval Sea Systems Command
NAVSEA 61R
Washington, D. C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 81323
Washington, D. C. 20360

Aircrew Systems Branch
Systems Engineering Test Directorate
U.S. Naval Test Center
Patuxent River, MD 20670

Mr. Milon Essoglou
Naval Facilities Engineering Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

CAPT Robert Biersner
Naval Biodynamics Laboratory
Michoud Station
Box 29407
New Orleans, LA 70189

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD

Dr. George Moeller
Human Factors Engineering Branch
Naval Submarine Base
Submarine Medical Research Lab.
Groton, CT 06340

Department of the Navy

Head
Aerospace Psychology Department
Naval Aerospace Medical Research Lab
Pensacola, FL 32508

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Dr. Jerry Tobias
Auditory Research Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Dr. Robert Blanchard
Code 17
Navy Personnel Research and
Development Center
San Diego, CA 92152-6800

LCDR T. Singer
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Stephen Merriman
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

LT. Dennis McBride
Human Factors Branch
Pacific Missle Test Center
Point Mugu, CA 93042

Dr. Kenneth L. Davis
Code 414
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

LCDR R. Carter
Office of Chief on Naval Operations
(OP-01B)
Washington, D.C. 20350

Dean of the Academic Departments
U. S. Naval Academy
Annapolis, MD 21402

CDR W. Moroney
Naval Air Development Center
Code 602
Warminster, PA 18974

Human Factor Engineering Branch
Naval Ship Research and Development
Center, Annapolis Division
Annapolis, MD 21402

Dr. Harry Crisp
Code N 51
Combat Systems Department
Naval Surface Weapons Center
Dahlgren, VA 22448

Mr. John Quirk
Naval Coastal Systems Laboratory
Code 712
Panama City, FL 32401

Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Dr. Charles Holland
Office of Naval Research Branch Office
London
Box 39
EPO New York 09510

Dr. Rabinder N. Madan
Code 414
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Eugene E. Gloye
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

Dr. David Mizell
ONR Detachment
1030 Green Street
Pasadena, CA 91106-2485

Dr. Glen Allgaier
Artificial Intelligence Branch
Code 444
Naval Electronics Ocean System Center
San Diego, CA 92152

Dr. Steve Sacks
Naval Electronics Systems Command
Code 61R
Washington, D.C. 20363-5100

Dr. Sherman Gee
Command and Control Technology,(MAT 0721)
Office of Naval Technology,
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Robert A. Fleming
Human Factors Support Group
Naval Personnel Research & Development Ctr.
1411 South Fern Street
Arlington, VA 22202

Dr. Dick Kelly
Human Factors Division, Code 17
Naval Personnel Research & Development Center
San Diego, CA 92152-6800

Department of the Army

Dr. Edgar M. Johnson
Technical Director
U.S. Army Research Institute
Alexandria, VA 22333-5600

Technical Director
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, MD 21005

Director, Organizations and Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Robert M. Sasmor
Director, Basic Research
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Department of the Air Force

Dr. Kenneth R. Boff
AF AMRL/HE
Wright-Patterson AFB, OH 45433

Dr. A. Fregly
U. S. Air Force Office of Scientific Research
Life Science Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332-6448

Mr. Charles Bates, Director
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks Air Force Base, TX 78235

Dr. R. K. Dismukes
Associate Director for Life Sciences
AFSOR
Bolling AFB
Washington, D.C. 20032-6448

Foreign Addresses

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF England

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Tech. Estab.
Teddington, Middlesex
TW11 OLN
England

Other Government Agencies

Dr. M. C. Montemerlo
Information Sciences &
Human Factors Code RC
NASA HQS
Washington, D.C. 20546

Dr. Alan Leshner
Deputy Division Director
Division of Behavioral and
Neural Sciences
National Science Foundation
1800 G. Street, N.W.
Washington, D.C. 20550

Defense Technical Information
Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 copies)

Dr. Clinton Kelly
Defense Advanced Research
Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Other Organizations

Dr. Harry Snyder
Dept. of Industrial Engineering
Virginia Polytechnic Institute
and State University
Blacksburg, VA 24061

Dr. Amos Tversky
Dept. of Psychology
Stanford University
Stanford, CA 94305

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311

Dr. J. O. Chinnis, Jr.
Decision Science Consortium,
Inc.
7700 Leesburg Pike
Suite 421
Falls Church, VA 22043

Dr. T. B. Sheridan
Dept. of Mechanical
Engineering
Massachusetts Institute of
Technology
Cambridge, MA 02139

Dr. Daniel Kahneman
The University of British
Department of Psychology
#154-2053 Main Mall
Vancouver, British Columbia
Canada V6T 1Y7

Dr. Stanley Deutsch
NAS-National Research Council
(COHF)
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Dr. Meredith P. Crawford
American Psychological
Association
Office of Educational Affairs
1200 17th Street N.W.
Washington, D.C. 20036

Dr. Deborah Boehm-Davis
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Dr. Paul E. Lehner
PAR Technology Corp.
7926 Jones Branch Drive
Suite 170
McLean, VA 22102

Other Organizations

Dr. Babur M. Pulat
Department of Industrial Engineering
North Carolina A&T State University
Greensboro, NC 27411

Dr. Stanley N. Roscoe
University of Colorado
Boulder, CO 80309

Dr. James H. Howard, Jr.
Department of Psychology
Catholic University
Washington, D. C. 20064

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801

Dr. Robert Wherry
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics Co.
St. Louis Division
Box 516
St. Louis, MO 63166

Dr. Lola L. Lopes
Department of Psychology
University of Wisconsin
Madison, WI 53706

Dr. Stanley N. Roscoe
New Mexico State University
Box 5095
Las Cruces, NM 88003

Mr. Joseph G. Whol
Alphatech, Inc.
3 New England Executive Park
Burlington, MA 10803

Dr. Marvin Cohen
Decision Science Consortium, Inc.
Suite 721
7700 Leesburg Pike
Falls Church, VA 22043

Dr. William R. Utal
NOSC-Hawaii
Box 997
Kailua, HI 96734

Dr. William B. Rouse
School of Industrial and Systems Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Ms. Denise Benel
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Dr. Andrew P. Sage
Assoc. V. P. for Academic Affairs
George Mason University
4400 University Drive
Fairfax, VA 22030

Other Organizations

Dr. Richard Pew
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Hillel Einhorn
Graduate School of Business
University of Chicago
1101 E. 58th Street
Chicago, IL 60637

Dr. Douglas Towne
University of Southern California
Behavioral Technology Lab
1845 South Elena Avenue, Fourth Floor
Redondo Beach, CA 90277

Dr. James P. Jenkins
Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. John Payne
Graduate School of Business
Administration
Duke University
Durham, NC 27706

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Azad Madni
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Tomaso Poggio
Massachusetts Institute of Tech.
Center for Biological Information
Processing
Cambridge, MA 02139

Dr. Robert A. Hummel
New York University
Courant Inst. of Mathematical
Sciences
251 Mercer Street
New York, New York 10012

Dr. H. McI. Parsons
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Dr. Paul Solvic
Decision Research
1201 Oak Street
Eugene, OR 97401

Other Organizations

Dr. David Castanon
ALPHATECH, Inc.
111 Middlesex Turnpike
Burlington, MA 01803

Dr. E. Douglas Jensen
Carnegie-Mellon University
Computer Science Dept.
Pittsburgh, PA 15213

Dr. David Noble
Engineering Research Assoc.
8616 Westwood Center Drive
McLean, VA 22180

Dr. Bruce Hamill
The Johns Hopkins Univ.
Applied Physics Lab
Laurel, MD 20707

Dr. A. Ephremides
University of Maryland
Electrical Engineering Dept.
College Park, MD 20742

Dr. Michael Athans
Massachusetts Inst. of Technology
Lab Information & Decision Systems
Cambridge, MA 01803

Dr. Leonard Adelman
PAR Technology Corp.
7962 Jones Branch Drive
Suite 170
McLean, VA 22102

Dr. Baruch Fischhoff
Perceptronics, Inc.
6271 Variel Ave.
Woodland Hills, CA 91367

END

FILMED

3-86

DTIC